? b 155,5,357
29apr03 07:19:24 User208669 Session D2272.1
$0.32 0.092 DialUnits File1
$0.32 Estimated cost File1
$0.03 TELNET
$0.35 Estimated cost this search
$0.35 Estimated total session cost 0.092 DialUnits

SYSTEM:OS - DIALOG OneSearch
File 155:MEDLINE(R) 1966-2003/Apr W3
(c) format only 2003 The Dialog Corp.
*File 155: Medline has been reloaded and accession numbers have changed. Please see HELP NEWS 155.
File 5:Biosis Previews(R) 1969-2003/Apr W3
(c) 2003 BIOSIS
*File 5: Alert feature enhanced for multiple files, duplicates removal, customized scheduling. See HELP ALERT.
File 357:Derwent Biotech Res. _1982-2003/Apr W4
(c) 2003 Thomson Derwent & ISI
*File 357: File is now current. See HELP NEWS 357.
Alert feature enhanced for multiple files, etc. See HELP ALERT.

Set Items Description
--- ----- -----------
? ds
Set Items Description
S1 85564 GC OR CG OR G(W)C OR C(W)G
S2 755771 RNA
S3 8346 S1 AND S2
S4 1110894 VIRUS OR PHAGE OR VIRUSES OR PHAGES
S5 52004 BACTERIOPHAGE OR BACTERIOPHAGES
S6 8346 S1 AND S2 AND (S3 OR S4)
S7 5872 S1(W)CONTENT
S8 1551 S6 AND S7
S9 284592 VECTOR OR VECTORS
S10 55 S8 AND S9
S11 48 RD (unique items)
S12 52951 S2(3N)(S4 OR S5)
S13 40 S12 AND S7
S14 36 S13 NOT S10
S15 24 RD (unique items)
S16 44782 G(2W)C OR C(2W)G
S17 21971 S16 NOT S1
S18 1027 S17 AND S2 AND (S4 OR S5)
S19 4109 S16(W)CONTENT
S20 2 S19 AND S18
S21 6 S16 AND CONTENT AND S18
S22 121016 OPTIMI?
S23 129 S1 (3N)S22
S24 3 S2 AND S23
S25 1 S23 AND (S4 OR S5)
S26 6409 S1(3N)CONTENT
S27 1627 S2 AND S26
S28 156 S27 AND (S4 OR S5)
S29 52951 S2(3N)(S4 OR S5)
S30 40 S28 AND S29
S31 27 RD (unique items)
S32 24 S31 NOT S11
S33 26825 POLIO?
S34 8 S26 AND S33
S35 5628782 AT OR TA OR A(2W)T OR T(2W)A
S36 5694 S35 AND S2 AND (S4 OR S5) AND S9
S37 18393 S35(3N) CONTENT
S38 10 S36 AND S37
S39 10 RD (unique items)
? t s11/7/1 6 7 13 16 22 48
11/7/1 (Item 1 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 2003 The Dialog Corp. All rts. reserv.
13956215 22216965 PMID: 12230944
Cloning and high level nonfusion expression of recombinant human basic fibroblast growth factor in Escherichia coli.
Chen Xiao-Jia; Sun Fen-Yong; Xie Qiu-Ling; Liao Mei-De; Zhang Ling; Li Zhi-Ying; Hong An; Lin Jian; et al
Bioengineering Institute of Jinan University, Guangzhou 510632, China. ojds@jnu.edu.cn.
Acta pharmacologica Sinica (China) Sep 2002, 23 (9) p782-6, ISSN 1671-4083 Journal Code: 100956087
Document type: Journal Article
Languages: ENGLISH
Main Citation Owner: NLM
Record type: In Process
AIM: To obtain high-level expression of nonfusion recombinant human basic fibroblast growth factor (rhbFGF). METHODS: hbFGF cDNA was prepared from the total RNA of embryonic brain tissue. As a template, the obtained gene was used to clone nonfusion rhbFGF. New primers were employed to alter the translation initiation region (TIR) and reduce the G+C content through nucleotide change. Using pET-3C as vector, the cloned rhbFGF was expressed in BL21 (DE3). RESULTS: rhbFGF was expressed in E coli up to 30 % of the total cellular protein. Cation exchange and heparin affinity chromatography were employed to purify the target protein from the supernatant of bacteria lysate. The bioactivity of the purified rhbFGF was identical with the standard bFGF. CONCLUSION: Modification of TIR is an effective means to increase nonfusion expression rate of recombinant proteins, such as rhbFGF,

in E coli.
Record Date Created: 20020916

11/7/6 (Item 6 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 2003 The Dialog Corp. All rts. reserv.
09859106 21671910 PMID: 11812218
Optimization of the expression of equistatin in Pichia pastoris.
Outchkourov Nikolay S; Stiekema Willem J; Jongsma Maarten A
Business Unit Cell Cybernetics, Plant Research International, Wageningen, The Netherlands.
Protein expression and purification (United States) Feb 2002, 24 (1) p18-24, ISSN 1046-5928 Journal Code: 9101496
Document type: Journal Article
Languages: ENGLISH
Main Citation Owner: NLM
Record type: Completed
To improve the expression of equistatin, a proteinase inhibitor from the sea anemone Actinia equina, in the yeast Pichia pastoris, we prepared gene variants with yeast-preferred codon usage and lower repetitive AT and GC content. The full gene optimization approximately doubled the level of steady-state mRNA and protein accumulated in the culture medium. The removal of a short stretch of 12 additional nucleotides from the multiple cloning site (MCS) sequence in the vector pPIC9 had an enhancement effect similar to full gene optimization (factor 1.5) at the mRNA level. However, at the protein level, this increase was 4- to 10-fold. The optimized gene without the MCS sequence yielded 1.66 g/L active protein in a bioreactor and was purified by a new two-step procedure with a recovery of activity that was >95%. This production level constitutes an overall improvement of about 20-fold relative to our previously published results. The characteristics of the MCS sequence element are discussed in the light of its apparent ability to act as negative expression regulator. Copyright 2002 Elsevier Science (USA).
Record Date Created: 20020128
Record Date Completed: 20021003

11/7/7 (Item 7 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 2003 The Dialog Corp. All rts. reserv.
09467207 21240764 PMID: 11343134
Evolution of base composition and codon usage bias in the genus Flavivirus.
Jenkins G M; Pagel M; Gould E A; de A Zanotto P M; Holmes E C
Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.
Journal of molecular evolution (United States) Apr 2001, 52 (4) p383-90, ISSN 0022-2844 Journal Code: 0360051
Document type: Journal Article
Languages: ENGLISH
Main Citation Owner: NLM
Record type: Completed
The extent to which base composition and codon usage vary among RNA viruses, and the possible causes of this bias, is undetermined in most cases. A maximum-likelihood statistical method was used to test whether base composition and codon usage bias covary with arthropod association in the genus Flavivirus, a major source of disease in humans and animals. Flaviviruses are transmitted by mosquitoes, by ticks, or directly between vertebrate hosts. Those viruses associated with ticks were found to have a significantly lower G+C content than non-vector-borne flaviviruses and this difference was present throughout the genome at all amino acids and codon positions. In contrast, mosquito-borne viruses had an intermediate G+C content which was not significantly different from those of the other two groups. In addition, biases in dinucleotide and codon usage that were independent of base composition were detected in all flaviviruses, but these did not covary with arthropod association. However, the overall effect of these biases was slight, suggesting only weak selection at synonymous sites. A preliminary analysis of base composition, codon usage, and vector specificity in other RNA virus families also revealed a possible association between base composition and vector specificity, although with biases different from those seen in the Flavivirus genus.
Record Date Created: 20010508
Record Date Completed: 20010726

11/7/13 (Item 13 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 2003 The Dialog Corp. All rts. reserv.
06986585 91227136 PMID: 2027754
Expression of tetanus toxin fragment C in yeast: gene synthesis is required to eliminate fortuitous polyadenylation sites in AT-rich DNA.
Romanos M A; Makoff A J; Fairweather N F; Beesley K M; Slater D E; Rayment F B; Payne M M; Clare J J
Department of Molecular Biology, Wellcome Biotech, Beckenham, Kent, UK.
Nucleic acids research (ENGLAND) Apr 11 1991, 19 (7) p1461-7, ISSN 0305-1048 Journal Code: 0411011
Document type: Journal Article
Languages: ENGLISH
Main Citation Owner: NLM
Record type: Completed
Fragment C is a non-toxic 50 kDa fragment of tetanus toxin which is a candidate subunit vaccine against tetanus. The AT-rich Clostridium tetani DNA encoding fragment C could not be expressed in Saccharomyces cerevisiae due to the presence of several fortuitous polyadenylation sites which gave rise to truncated mRNAs. The polyadenylation sites were eliminated by chemically synthesising the DNA with increased GC-content (from 29% to

47%). Synthesis of the entire gene (1400 base pairs) was necessary to generate full-length transcripts and for protein production in yeast. Using a GAL1 promoter vector, fragment C was expressed to 2-3% of soluble cell protein. Fragment C could also be secreted using the alpha-factor leader peptide as a secretion signal. The protein was present at 5-10 mg/l in the culture medium in two forms: a high molecular mass hyper-glycosylated protein (75-200 kDa) and a core-glycosylated protein (65 kDa). Intracellular fragment C was as effective in vaccinating mice against tetanus authentic fragment C. The glycosylated material was inactive, though it was rendered fully active by de-glycosylation.
Record Date Created: 19910611
Record Date Completed: 19910611

11/7/16 (Item 16 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 2003 The Dialog Corp. All rts. reserv.
05929957 88284374 PMID: 2456256
Alteration of amino-terminal codons of human granulocyte-colony-stimulati ng factor increases expression levels and allows efficient processing by methionine aminopeptidase in Escherichia coli.
Devlin P E; Drummond R J; Toy P; Mark D F; Watt K W; Devlin J J
Department of Molecular Biology, Cetus Corporation, Emeryville, CA 94608.
Gene (NETHERLANDS) May 15 1988, 65 (1) p13-22, ISSN 0378-1119
Journal Code: 7706761
Document type: Journal Article
Languages: ENGLISH
Main Citation Owner: NLM
Record type: Completed
We have improved the expression of recombinant human granulocyte-colony-stimulating factor (G-CSF), produced by either pL or trpP expression vectors in Escherichia coli, by altering the sequence at the 5' end of the G-CSF-coding region. Initial attempts to express G-CSF resulted in neither detectable G-CSF mRNA nor protein in the trpP system, and only G-CSF mRNA was detectable in the pL system. We modified both expression vectors to decrease the G + C content of the 5' end of the coding region without altering the predicted amino acid sequence. This resulted in expression of detectable G-CSF mRNA and protein in both systems. Expression reached 17% and 6.5% of the total soluble cellular protein in the pL and trpP expression systems, respectively. The N-terminal sequence of the recombinant G-CSF from the pL system was Met-Thr-Pro-Leu-Gly-Pro-. G-CSF isolated from several human cell lines (including the LD-1 cell line reported here), does not have an N-terminal methionyl residue. Deletion of the threonine codon at the beginning of the coding region for the mature G-CSF resulted in efficient removal of the N-terminal methionine residue during expression in E. coli.
Record Date Created: 19880907
Record Date Completed: 19880907

11/7/22 (Item 4 from file: 5)
DIALOG(R)File 5:Biosis Previews(R)
(c) 2003 BIOSIS. All rts. reserv.
11609503 BIOSIS NO.: 199800391265
Contrasting expression of KAL in cell-free systems: 5' UTR and coding region structural effects on translation.
AUTHOR: de Zoysa Priyal A(a); Helliwell Rachel J A; Duke Veronique M; Quinton Richard; Bouloux Pierre-Marc G
AUTHOR ADDRESS: (a)Cent. Neuroendocrinol., Royal Free Hosp. Sch. Med., Room 2 609, Rowland Hill Street, London NW3 **UK
JOURNAL: Protein Expression and Purification 13 (2):p235-242 July, 1998.
ISSN: 1046-5928
DOCUMENT TYPE: Article
RECORD TYPE: Abstract
LANGUAGE: English
ABSTRACT: We investigated the expression of two different X-linked Kallmann (KAL) gene cDNAs in two different cell-free systems using rabbit reticulocyte lysate: (system A) transcription/translation coupled and (system B) noncoupled. System A yielded a single band of 76 kDa corresponding to anosmin-1, the expected full-length gene product, and upon addition of canine microsomal membranes produced a 85-kDa glycosylated form. System B did not produce any detectable protein band despite the expression of a beta-galactosidase-positive control gene. The first 179 bases of the coding sequence are 74% GC-rich and showed the potential to form imperfect hairpin structures, which in part may explain the translation inhibition of KAL in system B. This has further led us to speculate that coupling transcription to translation may either be preventing translating-inhibiting hairpin formation or be compensating for the lack of certain tissue-specific proteins in reticulocyte lysate that are essential in overcoming inhibitory hairpins during translation. Substitution of the 5'-UTR with an encephalomyocarditis virus internal ribosomal entry site (EMCV IRES) sequence resulted paradoxically in a lower yield of anosmin-1, suggesting that elements in the 5'UTR may be necessary for maintaining a "normal" level of expression. The use of KAL and luciferase reporters (containing different 5'UTRs) demonstrated that the native KAL 5' UTR is not involved in translational efficiency. However, this sequence may influence faithful translation initiation. Theoretical RNA conformation data imply that effective EMCV IRES usage with KAL may require favorable pairing between the IRES and unidentified sequences within the 5' coding region of the gene. This work provides a foundation both for the investigation of KAL regulation and for the characterization of its function.

11/7/48 (Item 19 from file: 357)
DIALOG(R)File 357:Derwent Biotech Res.
(c) 2003 Thomson Derwent & ISI. All rts. reserv.
0081403 DBR Accession No.: 88-12252

Cloning and sequencing of Pseudomonas genes encoding vanillate-demethylase - expression in Escherichia coli; DNA sequence
AUTHOR: Brunel F; Davison J
CORPORATE SOURCE: Unit of Molecular Biology, International Institute of Cellular and Molecular Pathology, 75 Avenue Hippocrate, 1200 Brussels, Belgium.
JOURNAL: J.Bacteriol. (170, 10, 4924-30) 1988
CODEN: JOBAAY
LANGUAGE: English
ABSTRACT: A Pseudomonas sp. ATCC 19151 gene bank was constructed in Escherichia coli S17-1 using plasmid pJRD203 as vector. The recombinant DNA molecules were packaged in phage lambda particles and were transferred by conjugation into a vanB5 mutant of Pseudomonas sp. which was unable to use vanillate as C-source. A recombinant was isolated containing a 2,598 bp fragment encoding vanillate-demethylase, which converts vanillic acid to protocatechuic acid. Complementation and insertional inactivation analysis showed that the fragment carried 2 genes (vanA and vanB) cotranscribed from a promoter upstream of vanA. The DNA sequence of the fragment was determined, and 2 open reading frames of 987 and 942 bp were present in the transcribed orientation. These had a very high G+C content in the third base of each codon, characteristic of Pseudomonas chromosomal genes. Expression of the genes in Escherichia coli with a phage T7 RNA-polymerase promoter system gave rise to 2 polypeptides of mol.wt. 36,000 and 33,000 which were identified by deletion analysis as products of vanA and vanB respectively. The vanB gene product was related to the ferredoxin family. (28 ref)
? t s15/7/5 11 13 15 16
15/7/5 (Item 5 from file: 155)
DIALOG(R)File 155:MEDLINE(R)

11157186 98033196 PMID: 9367373
Dinucleotide and stop codon frequencies in single-stranded RNA viruses.
Rima B K; McFerran N V
School of Biology and Biochemistry, The Queen's University of Belfast, UK. b.rima@qub.ac.uk
Journal of general virology (ENGLAND) Nov 1997, 78 (Pt 11) p2859-70, ISSN 0022-1317 Journal Code: 0077340
Document type: Journal Article
Languages: ENGLISH
Main Citation Owner: NLM
Record type: Completed
To identify potential selection pressures which lead to RNA sequence conservation, we examined the occurrence rates of dinucleotides in 64 single-stranded RNA virus genomes. These viruses may offer a particular insight into these pressures since their RNA-dependent RNA polymerases lack proofreading capability. This potentiates introduction of mutations into their genomes, yet unidentified selection processes conserve the genomes to a large degree. We report a strong inverse correlation between the C+G content and the occurrence of the CpG dinucleotide (r=0.71) in the RNA virus genomes, in contrast to earlier reports (Karlin et al., 1994, Journal of Virology 68, 2889-2897). We also detected significant suppression of UpA, correlating inversely with genomic U+A content. These suppressions are coupled with over-representation of the complementary pair of dinucleotides, CpA and UpG. In addition, we highlight the fact that odds ratios for dinucleotides are not independent variables, a situation apparently not widely appreciated in the literature. This led us to view the over-representation of CpA and UpG as a consequential outcome of UpA and CpG suppression in the virus genomes. Potential factors influencing these disturbances are discussed. In addition, higher than random incidence was observed for 'out-of-frame' stop codons in the viral RNA genomes, with some preferences for individual codons being exhibited by certain virus groups. The UAG codon appeared more common in the +1 frame, the UGA in the -1 frame.
Record Date Created: 19971208
Record Date Completed: 19971208

15/7/11 (Item 11 from file: 155)
DIALOG(R)File 155:MEDLINE(R)

06655896 90281585 PMID: 2353453
Sequence of the genome RNA of rubella virus: evidence for genetic rearrangement during togavirus evolution.
Dominguez G; Wang C Y; Frey T K
Department of Biology, Georgia State University, Atlanta 30302-4010.
Virology (UNITED STATES) Jul 1990, 177 (1) p225-38, ISSN 0042-6822 Journal Code: 0110674
Contract/Grant No.: A1-00923; PHS; A1-21389; PHS
Document type: Journal Article
Languages: ENGLISH
Main Citation Owner: NLM
Record type: Completed
The nucleotide sequence of the rubella virus (RUB) genomic RNA was determined. The RUB genomic RNA is 9757 nucleotides in length [excluding the poly(A) tail] and has a G/C content of 69.5%, the highest of any RNA virus sequenced to data. The RUB genomic RNA contains two long open reading frames (ORFs), a 5'-proximal ORF of 6656 nucleotides and a 3'-proximal ORF of 3189 nucleotides which encodes the structural proteins. Thus, the genomic organization of RUB is similar to that of alphaviruses, the other genus of the Togavirus family, and the 5'-proximal ORF of RUB therefore putatively codes for the nonstructural proteins. Sequences homologous to three regions of nucleotide sequence highly conserved among alphaviruses (a stem-and-loop structure at the 5' end of the genome, a 51-nucleotide conserved sequence near the 5' end of the genome, and a 20-nucleotide

conserved sequence at the subgenomic RNA start site) were found in the RUB genomic RNA. Amino acid sequence comparisons between the nonstructural ORF of RUB and alphaviruses revealed only one short (122 amino acids) region of significant homology, indicating that these viruses are only distantly related. This region of homology is located at the NH2 terminus of nsP3 in the alphavirus genome. The RUB nonstructural protein ORF contains two global amino acid motifs conserved in a large number of positive-polarity RNA viruses, a motif indicative of helicase activity and a motif indicative of replicase activity. The order of the helicase motif and the nsP3 homology region in the RUB genome is reversed with respect to the alphavirus genome indicating that a genetic rearrangement has occurred during the evolution of these viruses.
Record Date Created: 19900718
Record Date Completed: 19900718

15/7/13 (Item 13 from file: 155)
DIALOG(R)File 155:MEDLINE(R)

06168105 89183595 PMID: 2928109
Codon usage and secondary structure of MS2 phage RNA.
Bulmer M
Department of Statistics, Oxford, UK.
Nucleic acids research (ENGLAND) Mar 11 1989, 17 (5) p1839-43, ISSN 0305-1048 Journal Code: 0411011
Document type: Journal Article
Languages: ENGLISH
Main Citation Owner: NLM
Record type: Completed
MS2 is an RNA bacteriophage (3569 bases). The secondary structure of the RNA has been determined, and is known to play an important role in regulating translation. Paired regions of the genome have a higher G+C content than unpaired regions. It has been suggested that this reflects selection for high G+C content to encourage pairing, but a re-analysis of the data together with computer simulation suggest that it is an automatic consequence in any RNA sequence of the way it folds up to minimise its free energy. It has also been suggested that the three registers in which pairing can occur in a coding region are used differentially to optimise the use of the redundancy of the genetic code, but re-analysis of the data shows only weak statistical support for this hypothesis.
Record Date Created: 19890509
Record Date Completed: 19890509

15/7/15 (Item 15 from file: 155)
DIALOG(R)File 155:MEDLINE(R)

03394684 81085464 PMID: 6255897
Secondary structures of influenza and Sendai Virus RNAs.
Khristova M L; Sokolova M V; Yaroslavtseva N G; Busse T L; Kharitonenkov I G
Archives of virology (AUSTRIA) 1980, 66 (3) p241-53, ISSN 0304-8608 Journal Code: 7506870
Document type: Journal Article
Languages: ENGLISH
Main Citation Owner: NLM
Record type: Completed
The secondary structures of influenza and Sendai virus RNAs were investigated by thermal denaturation, circular dichroism and proflavine binding methods. In 0.1 M NaCl about 60% of the bases of both RNAs were involved in secondary structure. The melting temperatures (Tm) of both viral RNAs were linear functions of the logarithm of the sodium ion concentration in solution, but under all ionic conditions the melting temperatures of Sendai virus RNA were higher than those of influenza virus RNA. At all ionic strengths the melting range of Sendai virus RNA was less than influenza virus RNA, indicating that the helical regions in Sendai virus RNA were longer than those in influenza virus RNA. Although Sendai virus RNA had a higher thermal stability than influenza virus RNA, hyperchromicity and circular dichroism data showed that Sendai virus RNA had less G+C content (34%) within the double stranded regions than influenza virus RNA (48%). The binding isotherms of Sendai and influenza virus RNA-proflavine complexes were studied at different ionic strengths. The number of binding sites of proflavine with influenza virus RNA were significantly lower than those with Sendai virus RNA. These results demonstrate the essential difference between the secondary and tertiary structures of the RNAs under study.
Record Date Created: 19810219
Record Date Completed: 19810219

15/7/16 (Item 16 from file: 155)
DIALOG(R)File 155:MEDLINE(R)

03176640 80130744 PMID: 6243983
Helix-coil transitions in double-stranded viral RNA. Fine resolution melting and ionic strength dependence.
Steger G; Muller H; Riesner D
Biochimica et biophysica acta (NETHERLANDS) Feb 29 1980, 606 (2) p274-84, ISSN 0006-3002 Journal Code: 0217513
Document type: Journal Article
Languages: ENGLISH
Main Citation Owner: NLM
Record type: Completed
Helix-coil transitions of double-stranded RNA from reovirus and infectious bursal disease virus were measured optically in aqueous medium of different ionic strengths. In RNA from reovirus four transitions and in RNA from infectious bursal disease virus two transitions were resolved and

evaluated quantitatively. The ionic strength dependence of their midpoint temperatures dTm/dlog[Na+] were 13.5 +/- 0.3 degrees C for reovirus RNA and 14.9 +/-0.7 degrees C for infectious bursal disease virus RNA. The midpoint temperatures extrapolated to 1 M ionic strength were 102.9, 104.3, 105.6, and 108.8 +/- 0.3 degrees C for reovirus RNA, and 108.8 +/- 1.8 and 109.6 +/- 1.0 degrees C for infectious bursal disease virus RNA. The G + C content of the regions in reovirus RNA melting in the different transitions were determined from the spectrum of the hypochromicity. The quantitative interpretation of the data is carried out on the basis of the ion condensation theory. It is estimated for double-stranded RNA of 100% G + C, that dTm/dlog[na+] = 8.4 degrees C. The two-dimensional dependence Tm = Tm (ionic strength, G + C content) is given. The ionic strength dependence in different double-stranded RNAs is correlated to the spacing of the phosphate backbone, secondary structure, and tertiary structure.

Record Date Created: 19800523

Record Date Completed: 19800523

? t s25/7

25/7/1 (Item 1 from file: 357)

DIALOG(R)File 357:Derwent Biotech Res.

(c) 2003 Thomson Derwent & ISI. All rts. reserv.

0202095 DBR Accession No.: 96-12866

Selection and metabolic correction of normal and Gaucher patient CD34+ hematopoietic progenitors with a therapeutic and selectable retro virus - recombinant vector-mediated glucosylceramidase and mouse thermostable antigen gene transfer to hematopoietic stem cell for e.g. Gaucher disease gene therapy (conference abstract)

AUTHOR: Medin J A; Pawliuk R; Amiri M; Brooks C; Kluepfel-Stahl S; Migita M; Humphries R K; Karlsson S

CORPORATE AFFILIATE: Nat.Inst.Neurol.Disord.+Stroke-Bethesda Nat.Inst.Health-Bethesda Terry-Fox-Lab.Vancouver

CORPORATE SOURCE: Development and Metabolic Neurology Branch, NINDS, NIH, Bethesda, MD 20892, USA.

JOURNAL: J.Invest.Med. (44, 3, 279A) 1996

ISSN: 1081-5589 CODEN: 5684D

CONFERENCE PROCEEDINGS: Biomedicine '96: Medicine Research from Bench to Bedside, Washington, DC, 3-6 May, 1996.

LANGUAGE: English

ABSTRACT: A novel recombinant retro virus vector was constructed using a MFG retro virus backbone to optimize glucosylceramidase (GC, EC-3.2.1.45) and downstream mouse heat stable antigen (HSA) gene expression. A bicistronic cloning cassette with minimal intervening nucleotides was created and inserted into the MFG vector. Amphotropic producer cell clones were derived from this MGCsIH plasmid and had higher GC enzyme activity and HSA cell surface expression with equal or less copies of the integrated provirus than the GsEH producer cells previously used. Normal and Gaucher disease patient CD34+ cells were transduced with supernatant from both producers using conditions previously demonstrated to transduce the majority of cells. In side-by-side analyses, 8% of cells infected with the LGsEH virus were positive over background HSA staining, while a full 27% of CD34+ cells infected with the MGCsIH virus were positive. Further, the intensity of fluorescence staining of individual cells for HSA was a log brighter with the MGCsIH virus. Such a method may be used for e.g. Gaucher disease gene therapy. (0 ref)

? t s32/7/2 5 6 13

32/7/2 (Item 2 from file: 155)

DIALOG(R)File 155:MEDLINE(R)

(c) format only 2003 The Dialog Corp. All rts. reserv.

11367498 98248212 PMID: 9588798

Tetranucleotide frequencies in microbial genomes.

Noble P A; Citek R W; Ogunseitan O A

Belle W. Baruch Institute for Marine Biology and Coastal Research, University of South Carolina, Columbia 29208, USA. noble@biol.sc.edu

Electrophoresis (GERMANY) Apr 1998, 19 (4) p528-35, ISSN 0173-0835

Journal Code: 8204476

Document type: Journal Article

Languages: ENGLISH

Main Citation Owner: NLM

Record type: Completed

A computational strategy for determining the variability of long DNA sequences in microbial genomes is described. Composite portraits of bacterial genomes were obtained by computing tetranucleotide frequencies of sections of genomic DNA, converting the frequencies to color images and arranging the images according to their genetic position. The resulting images revealed that the tetranucleotide frequencies of genomic DNA sequences are highly conserved. Sections that were visibly different from those of the rest of the genome contained ribosomal RNA, bacteriophage, or undefined coding regions and had corresponding differences in the variances of tetranucleotide frequencies and GC content. Comparison of nine completely sequenced bacterial genomes showed that there was a nonlinear relationship between variances of the tetranucleotide frequencies and GC content, with the highest variances occurring in DNA sequences with low GC contents (less than 0.30 mol). High variances were also observed in DNA sequences having high GC contents (greater than 0.60 mol), but to a much lesser extent than DNA sequences having low GC contents. Differences in the tetranucleotide frequencies may be due to the mechanisms of intercellular genetic exchange and/or processes involved in maintaining intracellular genetic stability. Identification of sections that were different from those of the rest of the genome may provide information on the evolution and plasticity of bacterial genomes.

Record Date Created: 19980626

Record Date Completed: 19980626

32/7/5 (Item 5 from file: 155)

DIALOG(R)File 155:MEDLINE(R)
(c) format only 2003 The Dialog Corp. All rts. reserv.
11157186 98033196 PMID: 9367373
Dinucleotide and stop codon frequencies in single-stranded RNA viruses.
Rima B K; McFerran N V
School of Biology and Biochemistry, The Queen's University of Belfast, UK. b.rima@qub.ac.uk
Journal of general virology (ENGLAND) Nov 1997, 78 (Pt 11) p2859-70, ISSN 0022-1317 Journal Code: 0077340
Document type: Journal Article
Languages: ENGLISH
Main Citation Owner: NLM
Record type: Completed
To identify potential selection pressures which lead to RNA sequence conservation, we examined the occurrence rates of dinucleotides in 64 single-stranded RNA virus genomes. These viruses may offer a particular insight into these pressures since their RNA-dependent RNA polymerases lack proofreading capability. This potentiates introduction of mutations into their genomes, yet unidentified selection processes conserve the genomes to a large degree. We report a strong inverse correlation between the C+G content and the occurrence of the CpG dinucleotide (r=0.71) in the RNA virus genomes, in contrast to earlier reports (Karlin et al., 1994, Journal of Virology 68, 2889-2897). We also detected significant suppression of UpA, correlating inversely with genomic U+A content. These suppressions are coupled with over-representation of the complementary pair of dinucleotides, CpA and UpG. In addition, we highlight the fact that odds ratios for dinucleotides are not independent variables, a situation apparently not widely appreciated in the literature. This led us to view the over-representation of CpA and UpG as a consequential outcome of UpA and CpG suppression in the virus genomes. Potential factors influencing these disturbances are discussed. In addition, higher than random incidence was observed for 'out-of-frame' stop codons in the viral RNA genomes, with some preferences for individual codons being exhibited by certain virus groups. The UAG codon appeared more common in the +1 frame, the UGA in the -1 frame.
Record Date Created: 19971208
Record Date Completed: 19971208

32/7/6 (Item 6 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 2003 The Dialog Corp. All rts. reserv.
10280513 96082318 PMID: 7483819
The nucleotide sequence of the fusion protein gene of the peste des petits ruminants virus: the long untranslated region in the 5'-end of the F-protein gene of morbilliviruses seems to be specific to each virus.
Meyer G; Diallo A
CIRAD/EMVT, Maisons-Alfort, France.
Virus research (NETHERLANDS) Jun 1995, 37 (1) p23-35, ISSN 0168-1702 Journal Code: 8410979
Document type: Journal Article
Languages: ENGLISH
Main Citation Owner: NLM
Record type: Completed
cDNA corresponding to the fusion protein (F) gene of the vaccine strain of peste des petits ruminants virus (PPRV) was cloned and sequenced. The gene was 2321 nucleotides long excluding the poly(A) tail. As with other morbilliviruses, it had a long G/C rich stretch of about 525 nucleotides. There was no start codon before position 489 in the nucleotide sequence. From the 489th nucleotide to the 549th nucleotide, there were 4 ATG codons, two of which were in frame. The fourth ATG codon was in the best context to act as a start codon for encoding a protein which will be composed of 546 amino acids with a predicted molecular weight 59,310 Da. The comparison of the nucleic acid sequences of different morbillivirus F-protein genes revealed that the 5'-end sequence of the mRNA is specific to each virus. This sequence contains a long stretch of nucleotides rich in G/C content. When protein sequences were compared, it appears that, during evolution, substitutions in amino acid occurred in the F-protein of morbilliviruses such that the structure required for the fusion activity remains unchanged. These substitutions appear to have only occurred in the leader and the membrane anchor sequences, probably as an adaptation of the protein to the host cell.
Record Date Created: 19951212
Record Date Completed: 19951212

32/7/13 (Item 13 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 2003 The Dialog Corp. All rts. reserv.
06168105 89183595 PMID: 2928109
Codon usage and secondary structure of MS2 phage RNA.
Bulmer M
Department of Statistics, Oxford, UK.
Nucleic acids research (ENGLAND) Mar 11 1989, 17 (5) p1839-43, ISSN 0305-1048 Journal Code: 0411011
Document type: Journal Article
Languages: ENGLISH
Main Citation Owner: NLM
Record type: Completed
MS2 is an RNA bacteriophage (3569 bases). The secondary structure of the RNA has been determined, and is known to play an important role in regulating translation. Paired regions of the genome have a higher G+C content than unpaired regions. It has been suggested that this reflects selection for high G+C content to encourage pairing, but a re-analysis of the data together with computer simulation suggest that it is an automatic consequence in any RNA sequence of the way it folds up to minimise its free

energy. It has also been suggested that the three registers in which pairing can occur in a coding region are used differentially to optimise the use of the redundancy of the genetic code, but re-analysis of the data shows only weak statistical support for this hypothesis.
Record Date Created: 19890509
Record Date Completed: 19890509
? t s34/7/3
34/7/3 (Item 3 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 2003 The Dialog Corp. All rts. reserv.
09845949 21657440 PMID: 11799160
Novel design architecture for genetic stability of recombinant poliovirus: the manipulation of G/C contents and their distribution patterns increases the genetic stability of inserts in a poliovirus-based RPS-Vax vector system.
Lee Sang-Gu; Kim Dae-You; Hyun Byung-Hwa; Bae Yong-Soo
Department of Microbiology, Hannam University, Daejeon. Creagene Research Institute, Department of Genetic Engineering, Youngdong University, Youngdong. Korea Research Institute of Bioscience and Bioengineering, Daejeon South Korea.
Journal of virology (United States) Feb 2002, 76 (4) p1649-62, ISSN 0022-538X Journal Code: 0113724
Document type: Journal Article
Languages: ENGLISH
Main Citation Owner: NLM
Record type: Completed
Poliovirus has been studied as a live recombinant vaccine vector because of its attractive characteristics. The genetic instability, however, has hampered recombinant polioviruses (PVs) from being developed as an appropriate vaccine. A variety of different foreign inserts were cloned directly into our poliovirus Sabin 1-based RPS-Vax vector system, resulting in the production of recombinant PVs. The genetic stability of each recombinant PV was examined during 12 rounds of consecutive passage. It was found that the genetic stability of the recombinants was not well correlated with their insert size. Instead, elevated stability was frequently observed in recombinants with inserts of high G/C contents. Furthermore, a comparative study using different constructs of the human immunodeficiency virus env gene revealed that the internal deletion of the unstable insert was seemingly caused by the presence of the adjacent A/T-rich region. The instability of these inserts was completely remedied by (i) increasing the G/C contents and (ii) replacing the local A/T-rich region with the G/C-rich codon without a change of the amino acid. This means that stability is closely associated with the G/C content and the G/C distribution pattern. To see whether these findings can be applied to the design of genetically stable recombinant PV, we have reconstructed the heteromultimeric insert based on our design architecture, including the above-mentioned G/C rules and the template/ligation-free PCR protocol. The heteromultimeric insert was very unstable, as expected, but the manipulated insert with the same amino acid sequence showed complete genetic stability, not only in vitro, but also in vivo. Even though this guideline was established with our RPS-Vax vector system, to some extent, it can also be applied to other live viral vaccine vectors.
Record Date Created: 20020118
Record Date Completed: 20020222
? t s39/7/7 8
39/7/7 (Item 1 from file: 357)
DIALOG(R)File 357:Derwent Biotech Res.
(c) 2003 Thomson Derwent & ISI. All rts. reserv.
0304044 DBR Accession No.: 2003-05829 PATENT
Producing protein in recombinant expression system involves predicting secondary structure of RNA encoding a protein and increasing free energy for the secondary structure by modifying sequence of DNA encoding the RNA - vector-mediated gene transfer, expression in host cell and bioinformatic software for nucleic acid vaccine and gene therapy
AUTHOR: WEINER D B; YANG J
PATENT ASSIGNEE: WEINER D B; YANG J 2002
PATENT NUMBER: US 20020123099 PATENT DATE: 20020905 WPI ACCESSION NO.:
2003-066795 (200306)
PRIORITY APPLIC. NO.: US 971806 APPLIC. DATE: 20011004
NATIONAL APPLIC. NO.: US 971806 APPLIC. DATE: 20011004
LANGUAGE: English
ABSTRACT: DERWENT ABSTRACT: NOVELTY - Producing (M) a protein by translation of mRNA from heterologous DNA sequence (HDS), involves predicting the secondary structure of mRNA transcribed from a native HDS, modifying native HDS, where mRNA transcribed from modified HDS has a secondary structure with increased free energy compared to mRNA transcribed from native HDS and using modified HDS for protein production, is new. DETAILED DESCRIPTION - Producing (M) a protein in a recombinant expression system by translation of mRNA from heterologous DNA sequence (HDS), involves predicting the secondary structure of mRNA transcribed from a native HDS, modifying native HDS to produce modified HDS, where mRNA transcribed from modified HDS has a secondary structure with increased free energy compared to mRNA transcribed from native HDS and using modified HDS for protein production. INDEPENDENT CLAIMS are also included for the following: (1) an injectable pharmaceutical composition (PC) comprising a nucleic acid (NA) molecule (I) that includes a modified coding sequence encoding a protein operably linked to regulatory elements, where the modified coding sequence comprises a higher AT or AU content relative to the AT or AU content of the native coding sequence; and (2) a recombinant viral vector (II) comprising (I). BIOTECHNOLOGY - Preferred Method: The secondary structure of the mRNA transcribed from a native heterologous DNA sequence is predicted

using a computer and a computer program, and is modified by increasing the AT content of the coding sequence at the 5' end of the coding sequence, within 200 or 150 or 100 nucleotides from the initiation codon such that mRNA transcribed from it has an increased AU content. Preferred Composition: In PC, the modified coding sequence encodes an immunogen such as a pathogen derived proteins, a cancer antigen, autoimmune disease associated protein, or their immunogenic fragments, or a fusion protein that includes the above proteins or their immunogenic fragments, or encodes a non-immunogenic therapeutic protein such as cytokines, growth factors, blood products or enzymes. The modified coding sequence comprises dispersed modifications which are up to 100, 150 or 200 bases in length alternating with regions of native coding sequence. The sequence comprises a higher AT or AU content in the first 100, 150 or 200 bases relative to the AT or AU content of the native nucleic acid sequence. ACTIVITY - Immunosuppressive; Antirheumatic; Cytostatic; Neuroprotective; Antiarthritic; Antidiabetic; Antithyroid; Antiulcer; Antipsoriatic; Virucide; Antiparasitic; Antiallergic. No biological data given. MECHANISM OF ACTION - Gene therapy. USE - (M) is useful for producing a protein in a recombinant expression system, preferably a cell free in vitro transcription and translation system, an in vitro cell expression system, a DNA construct used in direct DNA injection, or a recombinant vector for delivery of DNA to an individual (claimed). PC is useful for eliciting broad immune responses against a target protein, i.e. proteins specifically associated with pathogens such as viruses, parasites, allergens, or the individual's own abnormal cells. PC confers a broad based protective immune response against hyperproliferative cells that are characteristic in hyperproliferative diseases including all forms of cancer and psoriasis. PC is also useful for treating individuals suffering from autoimmune diseases including rheumatoid arthritis, multiple sclerosis, Sjogren's syndrome, insulin dependent diabetes mellitus, autoimmune thyroiditis, Crohn's disease, ulcerative colitis, and psoriasis. EXAMPLE - The addition of a leader sequence to minimize free energy in the West Nile virus Capsid mRNA resulted in enhanced protein expression and immune response. To enhance the transcription and translation efficiency of transgenes, the human IgE leader sequence was added to the 5' upstream of open reading frame (orf) sequences. The addition of a sequence encoding the human IgE leader sequence containing codons that were less prevalently utilized in humans (WNBy-DJY construct (yeast codon)) resulted in a predicted secondary structure for the mRNA with an increased free energy value, relative to the secondary structure for the mRNA without the leader sequence (WNVwt construct (wild type)), or relative to the secondary structure for the mRNA encoding a leader sequence optimized with human codons (WNVh-DJY construct (human codon)). The construct encoding the leader sequence containing codons that were less prevalently utilized in humans (yeast optimized) yielded a higher level of protein than did the construct encoding the leader sequence containing human optimized codons, as determined by immunoprecipitation of radiolabeled in vitro translated proteins. The codons more prevalently used by yeast were, in general, AU rich and the codons more prevalently used by Homo sapiens were, in general, more GC rich. DNA plasmid injection into mouse muscle induced an antigen-specific, CD4+, Th cell-dependent immune response, as determined by intracellular interferon-gamma (IFN-gamma/flow cytometry analysis. The CD4+ Th cell-dependent, intracellular IFN-gamma production was quantitated by flow cytometry. Splenocytes isolated from pWNVy-DJY (pWNVCy)-immunized mice, expressed higher levels IFN-gamma, upon stimulation with in vitro translated Cp protein, than did the splenocytes isolated from pWNVh-DJY (pWNVCh)-immunized mice. (25 pages)

39/7/8 (Item 2 from file: 357)
DIALOG(R)File 357:Derwent Biotech Res.

0292665 DBR Accession No.: 2002-14512 PATENT
Producing recombinant protein for preparing pharmaceutical compounds to treat, e.g., cancers or autoimmune disorders, comprises predicting secondary structure (SS) of mRNA and modifying DNA to give mRNA with SS having increased free energy - virus vector-mediated gene transfer and expression in host cell for use in recombinant vaccine and nucleic acid vaccine preparation and hyperproliferative disease, rheumatoid arthritis, multiple sclerosis and diabetes mellitus therapy and gene therapy
AUTHOR: WEINER D B; YANG J
PATENT ASSIGNEE: UNIV PENNSYLVANIA 2002
PATENT NUMBER: WO 200229088 PATENT DATE: 20020411 WPI ACCESSION NO.:
2002-416682 (200244)
PRIORITY APPLIC. NO.: US 237885 APPLIC. DATE: 20001004
NATIONAL APPLIC. NO.: WO 2001US31451 APPLIC. DATE: 20011004
LANGUAGE: English
ABSTRACT: DERWENT ABSTRACT: NOVELTY - Producing (M1) a protein (I) in a recombinant expression system (II) comprises: (a) predicting the secondary structure of mRNA; (b) modifying the native heterologous DNA sequence where the mRNA transcribed from the modified DNA has a secondary structure with increased free energy; and (c) using the modified DNA in (II) for production of (I). DETAILED DESCRIPTION - Producing (M1) a protein (I) in a recombinant expression system (II) comprising translation of mRNA transcribed from a heterologous DNA sequence in (II), comprises: (a) predicting the secondary structure (SS) of mRNA transcribed from a native heterologous DNA sequence (III); (b) modifying (III) to produce a modified heterologous DNA sequence (IV) where mRNA transcribed from (IV) has a SS having increased free energy compared to the SS of the mRNA transcribed from the (III); and

(c) using (IV) in (II) for production of (I). INDEPENDENT CLAIMS are also included for the following: (1) an injectable pharmaceutical composition comprising a nucleic acid molecule that includes a modified coding sequence (IV) encoding a protein operably linked to regulatory elements, where (IV) comprises a higher AT or AU content relative to the AT or AU content of the native coding sequence and further comprising a pharmaceutical carrier; and (2) a recombinant viral vector comprising a nucleic acid molecule that includes (IV) encoding a protein operably linked to regulatory elements, where (IV) comprises a higher AT or AU content relative to the AT or AU content of the native coding sequence. BIOTECHNOLOGY - Preferred Method: In M1, the recombinant expression system is selected from a cell free in vitro transcription system, an in vitro cell expression system, a DNA construct used in direct DNA injection and a recombinant vector for delivery of DNA to an individual. The secondary structure of the mRNA transcribed from a native heterologous DNA sequence is predicated using a computer and computer program, and is modified by increasing the AT content of the coding sequence at the 5' end of the coding sequence such that mRNA transcribed has an increased AU content. The 5' end of the coding sequence is within 200 nucleotides, preferably 150 nucleotides, even more preferably, 100 nucleotides, from the initiation codon. Preferred Composition: (IV) has a higher AT or AU content in the first 200 bases, preferably 150 bases, even more preferably, 100 bases, relative to the AT or AU content of the native nucleic acid sequence. (IV) comprises a higher AT or AU content in at least one region of up to 200 bases, preferably 150 bases, even more preferably, 100 bases in length, relative to the AU or AT content of the native nucleic acid sequence. (IV) encodes an immunogen, such as a pathogen-derived proteins, immunogenic fragments, cancer antigen, immunogenic fragment, or a fusion protein encoding these. The immunogen can also be a fusion protein that includes an autoimmune disease-associated protein or immunogenic fragment. (IV) encodes a non-immunogenic therapeutic protein, such as cytokines, growth factors, blood products or enzymes. (IV) could also comprises dispersed modifications of up to 200 bases, preferably 150 bases more preferably, 100 bases in length alternating with regions of native coding sequence. (IV) comprises a higher AT or AU content in the last 200 bases, preferably 150 bases, more preferably, 100 bases relative to the AT or AU content of the native nucleic acid sequence. Preferred Vector: The vector comprises (IV) as above for the composition. ACTIVITY - Antirheumatic; antiarthritic; neuroprotective; immunosuppressive; dermatological; antidiabetic; antithyroid; antipsoriatic; vasotropic; antiallergic; gastrointestinal. No supporting data is given. MECHANISM OF ACTION - Vaccine; gene therapy. No supporting data is given. USE - For producing a protein in a recombinant expression system (claimed). Use of a nucleic acid or recombinant viral vector that have modified DNA sequences to improve protein production can be used in gene therapy and for the treatment of cancers, hyperproliferative diseases, and autoimmune diseases such as rheumatoid arthritis, multiple sclerosis, Sjogren's syndrome, sarcoidosis, insulin-dependent diabetes mellitus, autoimmune thyroiditis, reactive arthritis, ankylosing spondylitis, scleroderma, polymyositis, dermatomyositis, psoriasis, vasculitis, Crohn's disease and ulcerative colitis. ADMINISTRATION - Administration is topical, vaginal, rectal, intranasal, transdermal, oral, intravenous, subcutaneous, intraperitoneal, intramuscular, inhalation, insufflation, intrathecal or intraventricular. Dosage is 0.1 - 500 microg, preferably 100-200 microg DNA. ADVANTAGE - Predicting the secondary structure of mRNA to identify highly stable secondary structure that may interfere with ribosome function and reduce transcription can lead to improved protein expression levels. This can be brought about by modification of the native DNA sequence leading to expression of a modified mRNA sequence, e.g., increasing AU content in a coding sequence optimizes the sequence by reducing the corresponding RNA secondary structure's integrity and resulting in increased protein expression or translation by melting of the inhibitory secondary structures (stem loops) in the RNA transcripts.EXAMPLE - No suitable example is given.(47 pages)